

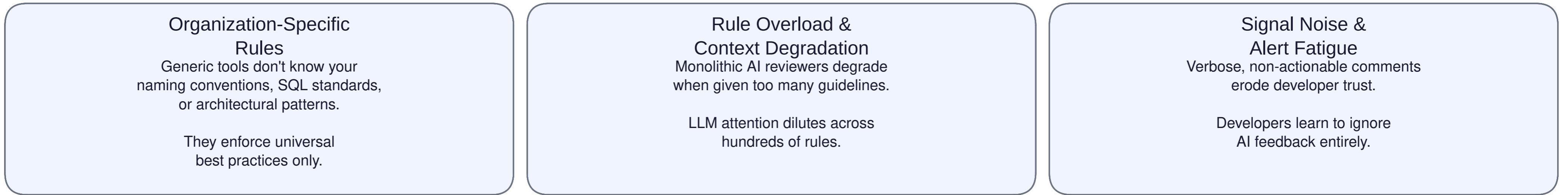
Filling the Gaps of AI Code Review in Enterprise

A Multi-Agent System for Organization-Specific Code Standards

Bc. Rastislav Budinský • Supervised by doc. Ing. Ivan Homoliak, Ph.D. • 2026

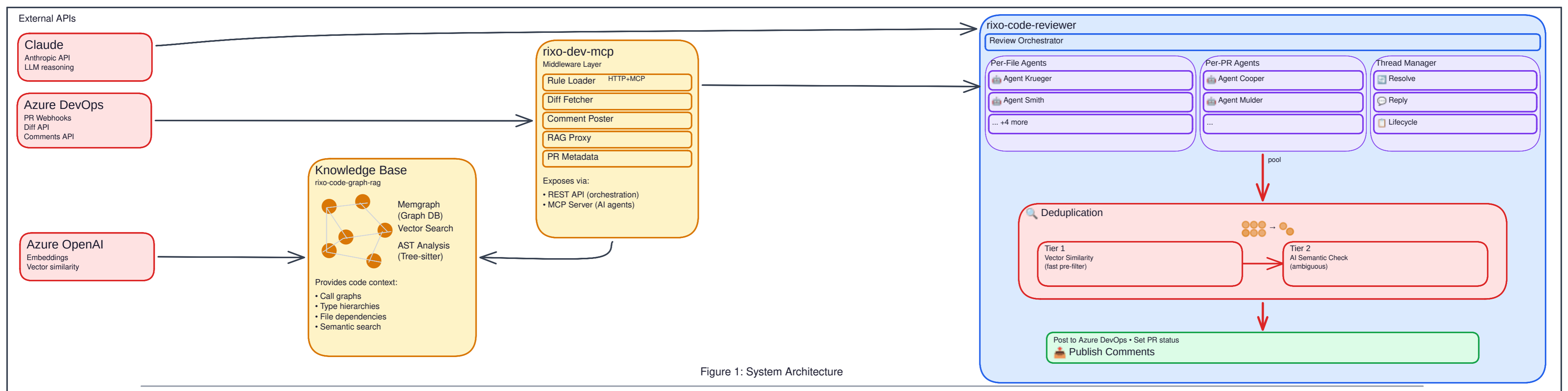
1 The Gap: What Commercial Tools Miss

Commercial AI code reviewers find bugs effectively — but they cannot enforce the specific coding standards your organization has built over years.



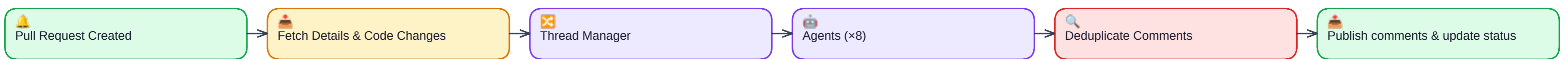
2 Our Approach: Specialized Multi-Agent Review

Instead of one monolithic reviewer, we split the task across specialized agents — each enforcing only its domain of rules.



3 Review Pipeline

— from webhook to published comment in a single automated flow



4 Results

— 436 pull requests, 5 months, production enterprise environment

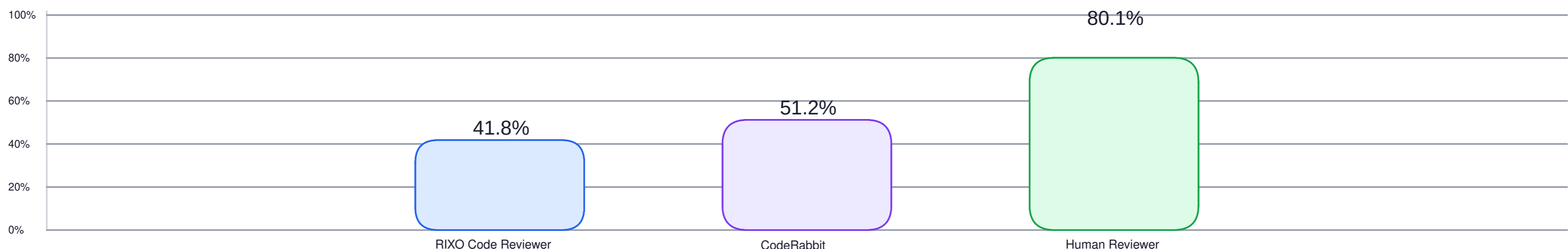
Focus	Human Design, logic	Rixio reviewer Style, rules	CodeRabbit Bugs, logic, flaws
Knows org rules?	✓	✓	✗
Cross-file bugs?	✓	✗	✓
Scales with volume?	✗	✓	✓
Data stays in-house?	✓	✓	✗
Uses AI buzzwords?	✗	Responsibly	Aggressively

Table 1: Complementary roles in code review

- 3x Coverage increase**
2.22 → 6.67 threads per pull request
- 41.8% Code change rate**
vs 51.2% CodeRabbit • 80.1% human
- 436 Pull requests**
5 months • production environment
- 8 Specialized agents**
style, security, SQL, testing, architecture...

5 Code Change Rate Comparison

— proportion of comments that led to actual code changes



Our system targets style & convention enforcement — a complementary aspect to bug-finding.

Figure 3: Code Change Rate by Reviewer Type

6 Lessons Learned

— from 5 months of production deployment

In Production — Components remain in active use at the company

Bc. Rastislav Budinský
xbudin05@stud.fit.vutbr.cz



⚠️ Deploy quality from day one
Premature deployment with incomplete deduplication eroded developer trust — a deficit that persisted even after significant quality improvements.

✓ Specialize, don't generalize
Eight focused agents outperform one monolithic reviewer. Each agent sees only rules relevant to its domain, reducing hallucination and noise.

✓ Complement, don't replace
AI style enforcement frees human reviewers to focus on design and logic — total coverage tripled while human effort remained stable.